2-2017

Novel gene-sequence markers for isolate tracking within *Monilinia fructicola* lesions

Madeline E. Dowling
Clemson University

Guido Schnabel
Clemson University, schnabe@clemson.edu

Harriet G. Boatwright
Clemson University

Sydney E. Everhart
University of Nebraska-Lincoln, everhart@unl.edu

Published in *Pest Management Science* 73 (2017), pp. 1822–1829. doi:10.1002/ps.4544

Submitted 5 October 2016; revised 19 December 2016; published 3 February 2017
Supporting information is presented following the References.

Novel gene-sequence markers for isolate tracking within *Monilinia fructicola* lesions

Madeline E. Dowling,[1] Guido Schnabel,[1] Harriet G. Boatwright,[1] and Sydney E. Everhart [2]

1 Department of Plant and Environmental Sciences, Clemson University, Clemson, SC, USA
2 Department of Plant Pathology, University of Nebraska, Lincoln, NE, USA

Corresponding author — S.E. Everhart, Department of Plant Pathology, University of Nebraska, Lincoln, NE 68588, USA; email everhart@unl.edu

Abstract

Background: *Monilinia fructicola* is a diverse pathogen of pome and stone fruits that causes severe economic losses each year. However, little is known about inoculum flow within or between orchards and pathogen establishment in an orchard, because few methods exist for detecting diversity or tracking isolates over time. SSR loci are an effective option, but may be confounded by a high degree of mutability and potential sensitivity to abiotic stress.

Results: Through transcriptome analysis, we identified novel markers *mrr1*, *DHFR* and *MfCYP01* and validated stability of these markers under fungicide stress in natural infection sites. Nucleotide variation within *mrr1*, *DHFR* and *MfCYP01* sequences differentiated isolates at all spatial scales: within the same infection site, between trees and between two farms. Sequenced regions were also effective for matching isolates collected from blossoms at the beginning of the season to progeny in cankers obtained at the end of the season.

Conclusions: Collectively, results show that *mrr1*, *DHFR* and *MfCYP01* are able to accurately differentiate *M. fructicola* isolates at the population level, can be used to track isolates over time, and are more stable than SSRs under external stresses. Either by themselves or combined with SSR markers, these gene-encoding regions are a much-needed tool for better understanding *M. fructicola* population dynamics.

Keywords: brown rot, SSR markers, epidemiology; *Prunus persica,* peach trees

1 Introduction

The development of reliable molecular markers is critical for understanding pathogen biology, particularly population genetics and epidemiology. Using multiple marker types in combination for within-species differentiation can strengthen the accuracy of results because mutation rates and mechanisms of polymorphism generation are different. Thus, differing marker types will likely detect a broader spectrum of variation than a single marker could.[1] Availability of multiple marker types for a pathogen of interest allows researchers to choose the most appropriate marker(s) for their studies, which may differ based on the scope of the study and research question being addressed. For example, in isolate tracking studies, which follow the population dynamics of microbial plant pathogens over time, the persistence or demise of genotypes within the population can be detected by repeatedly sampling a population and identifying isolates or strains using molecular markers. An assumption typically made in these studies is that mutation rates in marker regions are constant and unaffected by exogenous sources of stress, as demonstrated by studies testing marker stability under typical laboratory conditions.[2] However, it is known that certain regions of the genome, such as microsatellites, are more sensitive to stress, and external stresses may result in increased mutation rates.[3] For population studies, a genetic marker that has stability will produce more accurate results than markers more prone to stress-induced mutation. For these reasons, it is optimal to have multiple marker types available for a pathogen, and development of new markers with validated reliability is an important step towards accurate understanding of pathogen epidemiology.[4]

In the case of *Monilinia fructicola*, a causal agent of brown rot on pome and stone fruits, there are relatively few effective molecular makers capable of accurately differentiating fungal isolates within a field. Historically, vegetative compatibility groups (VCGs) and random amplified polymorphic DNA (RAPD) analyses were used, but either suffered from limited sensitivity (VCGs) or were not reproducible (RAPDs).[5–7] More recently, whole-genome sequencing and SNP analysis became available, but were not considered cost effective.[8,9] Without availability of a reference genome, the most effective and sensitive method for fingerprinting *M. fructicola* isolates is microsatellite or simple-sequence repeat (SSR) genotyping. Development of SSR markers has resulted in a number of studies on population dynamics and diversity of *M. fructicola*.[6,10,11] Although previous studies have

characterized fine-scale population dynamics of *M. fructicola* over time within individual trees,[12] none included resampling of isolates from the same infection lesions over a growing season. This lack of resampling assumes that an infection lesion will possess the same genotype from beginning to end of the growing season, but that may not be a reliable assumption given the high polymorphism and mutability of these loci.

SSR loci are present within all eukaryotic genomes and known to mutate via polymerase slippage at a rate of 10^{-2}–10^{-6} mutations per generation.[13] In contrast, gene-encoding regions are known to be more stable, with mutation rates 1000–10,000,000 times less frequent, approximately 10^{-9} in eukaryotes.[14] Because of a higher mutation rate in SSR loci, there is a potential for mutation in clonally propagated progeny isolates that would prevent them from being traced back to the correct progenitor isolate, which is more problematic within populations that have high genetic diversity and large clonal populations, such as *M. fructicola*. Some studies suggest that stressful conditions, such as sublethal fungicide or antibiotic exposure, may amplify mutation frequencies throughout genomes as a mechanism to enable more rapid emergence of a phenotype that allows pathogen survival and reproduction after exposure to a mutagen.[3,15,16] There is also mounting evidence that other stressors, including herbicide and temperature extremes, are related to SSR changes in fungi.[17] One study showed that ISSR loci in the plant pathogen *Alternaria alternata* changed when exposed *in vitro* to isothiocyanates used for disease control.[18] In *M. fructicola*, a recent study investigated the effect of *in vitro* sublethal fungicide exposure on SSR loci.[19,20] Isolates were grown on a gradient of the fungicide azoxystrobin (Abound Flowable; Syngenta Crop Protection, Greensboro, NC) and serially transferred from the 50–100% inhibition zone once a week for a total of 12 weeks, where each transfer was considered a generation. Comparison of the SSR profile of the progenitor isolate (never exposed to fungicide) to the fungicide-exposed isolate showed allelic changes at SSR loci after as few as ten generations.[21] It is unknown whether changes at gene regions would also be observed for these isolates.

Owing to the sensitive nature of isolate tracking, gene regions with lower mutational frequencies may be more accurate and less prone to mutate than SSR loci in response to exogenous stress, such as fungicide-induced stress. As no such markers exist for *M. fructicola*, the goal of our study was twofold: to develop novel gene markers with polymorphic intergenic regions that have greater stability than SSRs, and to cross-validate these markers by comparing stability and resolution of both marker types for isolates exposed to fungicide stress *in vitro* and under field conditions. Results of this study will be important because the new gene regions will give researchers another and, in some experimental conditions, more effective tool for tracking isolates over time, as well as providing a complement to SSR genotyping. More broadly, isolate tracking with these markers enables a better understanding of the spread of inoculum and relative importance of sexual and asexual reproduction to disease development and *M. fructicola* spread in stone fruit orchards, as shown in our recent publication.[21]

2 Materials and Methods

2.1 Marker development for variable gene regions

Transcriptomes of four *M. fructicola* isolates from a previous study (Li X, unpublished) were used to identify variable gene regions for candidate marker development. Briefly, actively growing mycelia of four *M. fructicola* isolates were grown in triplicate and used to purify RNA, which was used for library preparation by the sequencing facility. Sequencing was performed on the Illumina HiSeq 2000 platform to yield 2.17 million 100 bp paired-end reads passing filter and obtained as a FASTQ file of raw data from the sequencing facility. Prior to assembly, raw read quality was assessed using FastQC v.0.10.1 and trimmed in a content-dependent method using ConDeTri v2.2, resulting in an average per-read PHRED quality score of 37. *De novo* assembly of each transcriptome was performed using Trinity, with 25 as the optimal *k*-mer length. This yielded a grand total of 38,272 contigs, with 20,301 unique, and 17,971 were alternate splice forms of these contigs (Li X, private communication, 2015). Outputs of Trinity were converted to sequence alignments containing genomic coordinates (BAM files) using RSEM v.1.2.8,[22] which allowed for visualization using IGV (Integrative Genome Viewer 2.3.26).[23] Only portions of the transcriptome with data from each of the four *M. fructicola* isolates were used for random selection and manually compared for nucleotide sequence polymorphisms. Gene sequence alignments of the four isolates were used to identify gene regions with a total of ten or more variable regions, including SNPs and insertions/deletions greater than or equal to 3 bp in length and no longer than 30 bp. The consensus sequence was used to determine the start position of each variant, with each allele present in as few as one in four isolates. *DHFR*, *MfCYP01* and *HM2* homologous genes in *M. fructicola* were the first regions detected that fit the criteria above. More regions would have been mined if isolate genotype resolution was insufficient. In addition to these genes, we also searched for the *mrr1* homologous gene in *M. fructicola* using a Basic Local Alignment Search Tool (BLAST) search of the 20,301 unique contigs, as we had previously observed intraspecific variability in sequencing results of this gene (Hu M, private communication, 2015). Primers were designed to anneal near the beginning and end of each gene sequence obtained from the consensus sequence of the transcriptome analysis. Proposed primer sequences were checked for melting temperature and secondary structure using NCBI Primer BLAST.

Sequences of candidate gene regions were obtained for five *M. fructicola* isolates originally collected from different cankers and blossoms as described below. Gene regions consistently showing more than five nucleotide differences between isolates were retained in the study; these regions were *mrr1*, *DHFR* and *MfCYP01*. Primers were designed to amplify the most variable portions of each gene. The longest region was *mrr1*, which was sequenced using external and internal reverse and forward primers (Table 1). Amplification of *DHFR* and *MfCYP01* did not require internal primers (Table 1). For *mrr1*, the PCR program began with an initial denaturation of 94 °C for 5 min, then 40 cycles of 94 °C for 30 s, 50 °C for 1 min and 72 °C for 2 min, and then a final extension of 72 °C for 10 min. The PCR program to amplify the *DHFR* fragment included an initial denaturation cycle of 95 °C for 3 min, followed by a cycle repeated 34 times of 95 °C for 30 s, 57 °C for 30 s and 72 °C for 1 min. Lastly, there was an extension cycle at 72 °C for 5 min. The program for amplifying the *MfCYP01* region began with a denaturation cycle of 94 °C for 3 min, then a cycle repeated 35 times of 94 °C for 30 s, 52 °C for 40 s and 72 °C for 2 min, and then a final extension cycle of 72 °C for 12 min. Sequencing was performed by CORE Laboratories at Arizona State University. Sequences were assembled and compared using Seqman and MegAlign, respectively, from the DNAstar Lasergene Suite (DNAstar Inc., Madison, WI).

Table 1. Primers used to amplify putative gene regions identified from the transcriptome of *M. fructicola*

Locus	Primer sequence (5′–3′)
DHFR	TGGCATATTTCGCGAGGGTT
	ATAGGAACGGGACCTGTGGA
mrr1 fragment 1	TGTCTTACCATCCTGTCATAC
	TCATTGTCTCAGTACCTCGATC
mrr1 fragment 2	TGCAGTCATGCGCTCTGATG
	TGTCTTACCATCCTGTCATAC
MfCYP01	TCACGAAACACACCCGTCT
	CCAGGCGTTATTGGCTATG

2.2 Field trials and fungicide applications

Two 11-year-old orchards established in 2003 and located in separate states were selected for experimental trials. The South Carolina trial contained 80 nectarine trees, *Prunus persica* var. *nucipersica* cv. 'June Princess' at Clemson University's Musser Fruit Research Farm, Seneca, South Carolina. The Georgia trial consisted of approximately 200 peach trees, *P. persica* cv. 'O'Henry', at the University of Georgia Horticulture Research Farm, Watkinsville, Georgia. These different hosts were selected because there is no known population subdivision of *M. fructicola* on different host subspecies, where both cultivars in the present study are considered disease susceptible. Herbicides and insecticides were applied according to standard commercial practice.[24] Independent treatments of azoxystrobin, formulated as Abound Flowable, and propiconazole, formulated as Tilt (Syngenta Crop Protection), were applied to different blocks with an airblast sprayer once each week from fruit set to commercial ripeness at half the rate of standard commercial practice to ensure pathogen survival. Isolates were collected from three treatment blocks in the Musser Farm trial. Abound Flowable azoxystrobin was applied to the first block, no fungicide was applied to the second and Tilt was applied to the third. Both fungicides were applied for 10 weeks at half the recommended field rate. Fungicide applications at the UGA Horticulture Research Farm were identical, except that the block treated with Tilt was not present. No fungicides were applied within the post-dormancy period to blossom blight sample collection, or in the preharvest interval.

2.3 Isolate collection and single spore isolations

Conidia were collected after petal fall in April 2014 from actively sporulating blighted blossoms. Sterile, individually wrapped cotton swabs (Thermo Fisher Scientific, Waltham, MA) were used to collect conidia from symptomatic flowers by gently swirling the cotton tip in newly formed spore masses. Swabs were replaced into the sterile cover, labeled and placed into freezer bags containing solid desiccant, and then refrigerated at 4 °C until isolation of the pathogen in the lab.

Collection of conidia from infected blossoms was done in a non-destructive manner in order to allow sampling of the infected woody lesion later in the season. However, not every blighted blossom results in an infection penetrating into the woody tissue and developing into a canker, so the number of blighted blossoms swabbed in the spring was necessarily greater than the number of cankers collected late season. Conidia from a total of 380 sporulating infected blossoms were sampled from 19 trees within each orchard block, where 20 conidial samples were obtained from each tree canopy. Before conidia from each blossom were sampled, two weatherproof tags printed with the isolate identification number were attached to the branch on either side of the blossom to guide canker collection later in the season. Selection criteria for blossom sampling included that they be distant from other infected blossoms on the same branch, on otherwise healthy branches, scattered throughout the tree canopy and attached to branches with a diameter of ≥1 cm. Nearby healthy blossoms were stripped off branches to prevent confusion.

Conidia of *M. fructicola* on cotton swabs remain viable in dry conditions at 4 °C for up to 1 year or more after collection. No more than 2 months after collection, swabs were removed from storage and individually tapped over water agar plates so that conidia landing on the plate could be allowed to germinate in the dark at 22 °C for 16 h. For each sample, a single germinating conidium was identified under a dissecting microscope, then isolated with a scalpel and placed on potato dextrose agar (PDA) in preparation for DNA extraction and permanent storage.

Twigs with cankers were collected in July 2014 from the tagged locations where conidia were previously collected from a blighted blossom. Cankers ambiguously labeled because of tag sliding or tag loss were eliminated. As sporulation from cankers is known to be sporadic, twigs with cankers were collected (4–5 cm in length and about 1 cm in diameter) in order to induce sporulation in the lab, under controlled conditions. Twigs containing a single canker were sterilized in 10% bleach solution for 1 min, rinsed for 1 min in ultrapure filtered water and then dried in a laminar flow hood on paper towels and placed in labeled petri dishes. Each dish contained 9.0 cm diameter filter paper No. 410 (VWR, Radnor, PA), which was saturated with ultrapure water to create a moist chamber. Petri dish lids were left closed for 3 days in light at 22 °C and then propped open to reduce humidity and stimulate production of conidia. Little chance of cross-contamination was present because airflow was minimal and plates were removed after sporulation was detected. After 1–5 weeks, spores emerged and were recovered from approximately 45% of the cankers. A total of 110 isolates were obtained from sporulating cankers at the University of Georgia (25) and the Musser Farm (85). Spores were collected using sterile swabs and then stored in a plastic bag with desiccant. Single spore isolations were performed as described above.

2.4 DNA extraction

Purification of DNA from mycelium was performed as previously described[25] using 'the quick and safe method' except that sample DNA was washed with chilled ethanol and then centrifuged at 13,400 rcf for 3 min. Ethanol supernatant was discarded, and residual ethanol was evaporated in a laminar flow hood until the pellet dried. Approximately 50 µL of TE buffer (10m*m* Tris and 0.1mM EDTA, pH 8.0) was added to pelleted DNA, and tubes were placed in a water bath at 45 °C for several hours to ensure complete dissolution of the DNA pellet. DNA was stored at −20 °C until later use.

2.5 Fragment size analysis and comparison

For SSR analysis, PCR was performed using primers and touchdown PCR described previously.[11] SSR loci used in the present experiment were SED, SEF, SEI, SEL, SEN, SEP and SEQ, and to limit production of stutter peaks as a result of poly-A tail formation during PCR, forward primers were pig-tailed on the 5′ end with GTTTT (Table 2). In the present study, the fluorophore FAM

Table 2. SSR primers used for isolate genotyping and tracking, modified for indirect fluorescent labeling with CAG-tag sequence (bold) on forward primer and GTTT 'pigtail' (underlined) on reverse primer for stutter peak reduction

Locus	Primer sequence (5′ –3′)
SEF	**CAGTCGGGCGTCATCA**GACTATAGAGTTTTCTACGGATGG GTTTTGTCTCTCAACTTTTAAATCAGCC
SEL	**CAGTCGGGCGTCATCA**GAGTATAACCAACCCAACGGC GTTTAGAGATGGAGTCAGGAGTGTTG
SEN	**CAGTCGGGCGTCATCA**TGCGTGTCATGTCGTCC GTTTCGAGGCTTAACTTCCGTGC
SEP	**CAGTCGGGCGTCATCA**TAGGCCACAGCTGATACCG GTTTATCAATTGGTTTGGGTCCTTG
SEQ	**CAGTCGGGCGTCATCA**GGAGGTGGATGGTGGGTAG GTTTGGCTGTGGGTTGAGTGAG
SED	**CAGTCGGGCGTCATCA**TTGGCATGGCATTTGGAGC GTTTCCATTTTATTCATATCCAACGCCC
SEI	**CAGTCGGGCGTCATCA**CTCAAGCGGTGGCTCAAAG GTTTAACCACCACGACCACGAC

was attached directly to primers instead of using the HEX-CAG complementary primer, where direct FAM attachment allowed faster throughput of samples. Base pairs corresponding to the CAG primer used for indirect fluorescent labeling were added to the FAM-labeled primer so that resulting fragment sizes would be similar to those described previously.[11] Thermal cycler conditions used for amplification were: an initial denaturation cycle of 95 °C for 2 min and 30 s, followed by a touchdown cycle repeated 20 times of 95 °C for 20 s, 60 °C for 20 s, with temperature decreasing by 0.5 °C every cycle, and 72 °C for 30 s. The third cycle of 95°C for 20 s, 50 °C for 20 s and 72 °C for 30 s was repeated 15 times. Fragment sizing was performed on the ABI 3730 by CORE Laboratories at Arizona State University. Fragment sizes were analyzed using GeneMapper 4.0 software (Life Technologies, Carlsbad, CA) and also visually assessed to ensure accurate size calling.

2.6 Mutability of marker regions

To test the stability of the SSR and gene markers, we obtained previously purified DNA from two isolates both before treatment and also after ten generations of exposure to fungicide azoxystrobin, known to cause oxidative stress.[19,20] SSR fragments and *mrr1*, *DHFR* and *MfCYP01* sequences of this DNA were analyzed as described above and examined for any changes in either gene region or SSR loci.

2.7 Data analysis

The R package *poppr* was used to produce a genotype accumulation curve to identify the optimal number and combination of SSR loci and gene regions for genotype resolution.[26] The same package was also used to calculate pairwise genetic distances between SSR genotypes using Bruvo's distance, which is a quantitative metric used for SSR markers that follow a stepwise model of mutation, and visualized in a minimum spanning network. Genetic regions used for population genetic studies must meet an assumption of neutrality, as regions under selection result in non-random distribution of alleles. To assess evidence of selection on the gene regions *MRR1*, *DHFR* and *MfCYP01*, Tajima's *D* statistic was calculated. Tajima's *D* compares the average number of pairwise differences in a population to the number of segregating sites.[27] A Tajima's *D* value of zero would indicate that variation in the form of single nucleotide polymorphisms (SNPs) within these genes is neutral and not under selection. Tajima's *D* values were calculated using DnaSP v.5 software.[28]

3 Results

Four gene regions with ten or more variants between each pairwise combination of four *M. fructicola* transcriptomes were selected for primer design. These regions selected using a BLAST search of the sequences obtained from the transcriptome putatively identified these regions as the multidrug resistance regulator (*mrr1*),[29] the dihydrofolate reductase region (*DHFR*), a cytochrome *P450* region (*MfCYP01*) and a histone modification region (*HM2*). When five isolates were sequenced using each of the four regions, the HM2 region had few polymorphisms, with only 1 or 2 variants among a preliminary assessment of five isolates, and was subsequently removed from the study. The other regions appeared polymorphic, with more than five variants between isolate combinations, including SNPs and insertions/deletions from 3 to 30 bp in length. All variants were converted to presence/ absence binary data for subsequent analyses.

Tajima's *D* was calculated for isolates from each gene region and treatment population to ensure regions were not under selection. This estimation was not possible for isolates obtained from the University of Georgia Horticulture Farm as there were only three samples in each population and at least four sequences are required for estimating Tajima's *D*. None of the estimated Tajima's *D* values for each gene was found to have a significant deviation from the neutral hypothesis ($P > 0.1$).

Pairs of isolates (some of which were exposed to low-dose fungicide treatments) were obtained from field sites in South Carolina (85) and Georgia (25). A total of 220 isolates were collected and genotyped, representing pairs of isolates from 110 infection sites: 110 isolates from blighted blossoms (BB) and 110 from twig cankers (TC) corresponding to the infection site of each of the blighted blossoms. Pairs of isolates (BB:TC) from each infection site were compared to detect incongruence of genotypes at gene regions and SSR loci. We chose a large sample size to ensure that isolates with matching BB:TC genotypes would be detected, as blighted blossoms and possibly cankers could be the result of infection by multiple isolate genotypes. Each of the three gene regions was sequenced separately because a difference at any one gene sequence was sufficient to determine that the isolate pair had non-matching BB:TC genotypes, which meant that sequencing all three genes was not necessary. Thus, a process of elimination was used to identify isolate pairs with non-matching gene sequences. For each pair of BB:TC isolates, gene sequences were first obtained for *mrr1* and lastly for *MfCYP01*. As a result, 110 BB:TC isolate pairs were sequenced at *mrr1*, showing that 52 BB:TC pairs had non-matching gene sequences. The remaining 58 BB:TC pairs were sequenced at *DHFR*, which eliminated 26 non-matching pairs. Finally, 32 BB:TC pairs were sequenced at *MfCYP01*, eliminating only one isolate and resulting in a final total of 31 pairs of isolates that had matching genotypes at all three gene region loci. Pairs of BB:TC isolates obtained from the same infection site with identical sequences at all three loci were considered genetically identical

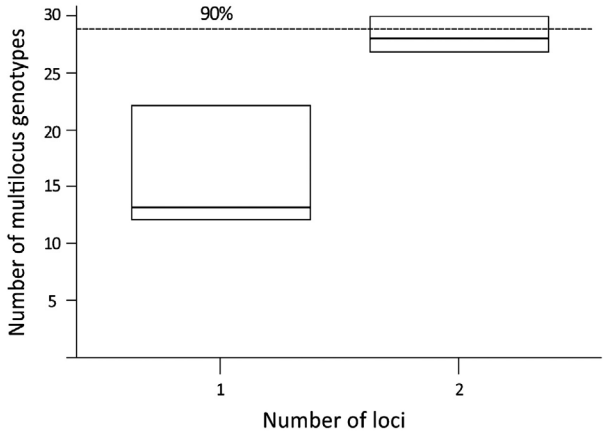

Figure 1. Number of multilocus genotypes resolved with increasing numbers of gene-encoding regions in a population of 31 isolates, where the number of genotypes resolved is estimated from all possible combinations of the three loci used in this study.

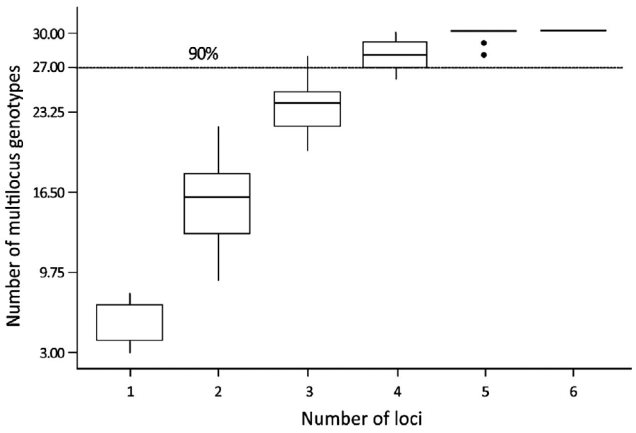

Figure 2. Number of multilocus genotypes resolved with increasing numbers of SSR loci in a population of 31 isolates, where the number of genotypes resolved is estimated from all possible combinations of the seven loci used in this study and approximately 90% of genotypes are resolved with four loci.

BB:TC pairs; genotypes are provided in supporting information Table S1. Thus, gene region analysis determined that 31 of 110 BB isolates had MLGs identical to the 31 corresponding TC isolates sampled from the same infection site later in the season, whereas 79 were non-matching within the same infection site.

Gene region sequences obtained for the 31 BB:TC pairs were compared between infection sites, with the identification of a total of 29 multilocus genotypes (93.5% resolution) when using all three loci (Fig. 1). The resolving power or ability of individual gene regions to differentiate multilocus genotypes between infection sites was estimated and was highest for *mrr1* (43.3%), moderate for *DHFR* (26.6%) and lowest for *MfCYP01* (16.6%). The combination of *DHFR* and *MfCYP01* resolved different MLGs for 70.0% of the isolates, while combinations *mrr1*+*DHFR* and *mrr1*+*MfCYP01* each resolved different MLGs for 73.3% of the isolates.

Resolution was not a direct product of the number of alleles detected at each locus, as *mrr1*, *DHFR* and *MfCYP01* had 21, 13, and 13 alleles respectively (Table 3). There was high variability between alleles at each locus, where average percentage nucleotide variability was 0.765% for *mrr1* (2280 bp long), 1.38% for *DHFR* (614 bp long) and 1.06% for *MfCYP01* (1658 bp long) alleles. The majority of differences between isolates were SNPs. However, there were multiple isolates containing whole codon deletions, and a few isolates with 20–30 bp long segments missing compared with all other isolates sequenced. Each missing codon or large section of DNA was considered a single mutational event. All gene regions contained both synonymous and non-synonymous nucleotide changes.

High variability existed between BB:TC isolate pairs with different MLGs. Between these non-matching isolates, percentage nucleotide variability, the average number of nucleotide differences per bp, was 1.1% for *mrr1* (2280 bp) and 1.8% for *DHFR* (614 bp). At the *MfCYP01* locus, only one isolate pair contained genetically different isolates, and the percentage variability between these isolates was 0.97% (1750 bp). Genetically distinct multilocus genotypes were differentiated by a minimum of four nucleotide polymorphisms.

To determine correspondence and relative resolution of *mrr1*, *DHFR* and *MfCYP01* gene loci as molecular marker types compared with SSR loci, fragment analysis was performed on the 31 pairs of BB:TC isolates that had identical gene region genotypes. SSR genotyping identified a total of 31 multilocus genotypes, where with gene regions 29 multilocus genotypes were observed between BB:TC pairs. Each locus was polymorphic, with an average of 4.1 alleles per locus, a minimum of two alleles at the locus SEL and a maximum of seven alleles at SEF. A genotype accumulation curve showed over 90% of multilocus SSR genotypes were resolved using four loci, and 100% of multilocus genotypes were resolved using five, six and seven loci (Fig. 2).

Overall, genotyping with seven SSR loci and three gene regions gave similar results, where BB:TC isolates from the same infection site had the same MLG. However, SSR analysis identified a total of 31 unique MLGs among all sites, where isolates from two sites designated as clones by the gene regions actually varied from each other at four SSR loci (SED, SEF, SEL and SEQ (Fig. 3). On average, these loci had more alleles than those that did not vary from pair to pair. The average number of loci for both farms for SED, SEF, SEL and SEQ was 4.5, while that for SEI, SEN and SEP was 3.7 alleles.

Table 3. Comparison of allele numbers and frequencies for SSR loci and gene regions

	SSR[a] loci							Gene regions		
	SED	SEF	SEI	SEL	SEN	SEP	SEQ	*mrr1*	*DHFR*	*MfCYP01*
Number of alleles	7	9	7	3	4	6	8	21	13	12
Frequency of MCG[b]	0.48	0.32	0.58	0.41	0.45	0.58	0.29	0.09	0.35	0.25
Frequency of LCG[c]	0.03	0.03	0.03	0.19	0.06	0.03	0.03	0.03	0.032	0.032

a. Simple sequence repeat loci.

b. Most common genotype.

c. Least common genotype.

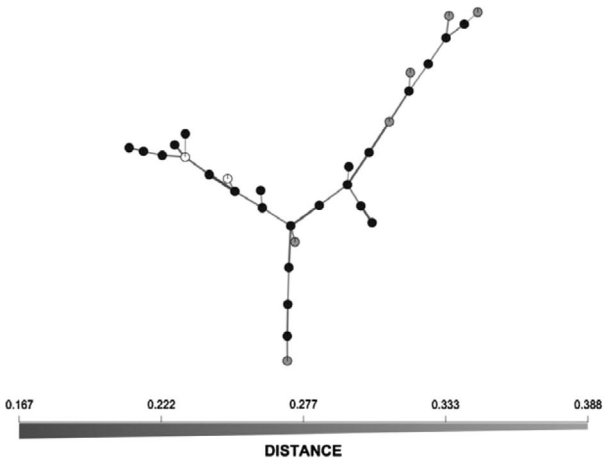

Figure 3. Genetic distance between SSR genotypes estimated using Bruvo's stepwise distance and visualized as a minimum spanning network. Each multilocus genotype is represented by a node (circle) and the genetic distance between each genotype is represented by edges (lines),where the thickness of the edge decreases with increasing genetic distance between genotypes and corresponds to the scale at the bottom of the figure. Black nodes and gray nodes represent Musser Farm and UGA BB:TC pairs, respectively, that differ from all other pairs at gene regions. White nodes represent the two Musser Farm BB:TC pairs that had identical gene region profiles.

To determine whether gene regions were more stable than SSR loci under fungicide-induced oxidative stress, seven-locus SSR multilocus genotypes and *mrr1*, *DHFR* and *MfCYP01* sequences were obtained for isolates before and after exposure to oxidative fungicide stress using DNA provided from a previous study.[19] Genotyped in the present study were two progenitor isolates and the two corresponding progeny isolates that had observable SSR mutations after sublethal fungicide exposure *in vitro*.[19,20] These isolates were sequenced in the present study at the gene regions *mrr1*, *DHFR* and *MfCYP01*. Although these isolates were shown to have mutations at five of the seven SSR loci, no changes were observed in gene sequences.

Overall, results obtained using gene loci and SSR loci were similar. Isolates considered genetically identical at the *mrr1*, *DHFR* and *MfCYP01* loci were verified as clones using SSR genotyping. Exceptions were two BB:TC isolate pairs from different lesions that had identical gene region sequences yet different SSR genotypes between lesions (isolates from the same lesion were not different). As expected, the main difference between SSR loci and gene regions was the mutation rates of these regions, where SSR loci are considered more polymorphic than intergenic regions. The two genotypes that differed at SSR regions but not at gene regions had fragment size differences at three loci: SEF, SEL and SEQ. These SSR loci followed a stepwise mutation model, with a single step difference in alleles at SEL, two steps different at SEF and four steps different at SEQ, resulting in a Bruvo's stepwise genetic distance of 160. Pairwise comparisons of these two genotypes with genotypes from other infection sites showed that these two were more closely related to each other (smaller Bruvo's genetic distance; obs=160) than they are to most other genotypes (Fig. 3). The average genetic distance of each of these genotypes to the Musser Farm isolates was more than twice as great (average genetic distance of 380.8 and 472.7) and 4 times as great compared with UGA farm isolates (average genetic distance of 666.0 and 717.2). Both genotypes were, however, individually more closely related to two other Musser Farm genotypes (Bruvo's distance<160), with as few as two loci differentiating those genotypes.

4 Discussion

The purpose of this study was to develop stable population genetic markers with power to differentiate *M. fructicola* isolates at the population level for use in isolate tracking experiments that sample the pathogen from the same infection site multiple times within a season. Sequenced segments of *mrr1*, *DHFR* and *MfCYP01* gene regions were shown to contain variable regions, and nucleotide BLAST of these sequences revealed their identities. Selection on these particular genes has not been documented in *M. fructicola*. However, the *mrr1* region is related to DMI resistance in other fungi, and genes in the cytochrome P450 superfamily are related to DMI resistance.[30,31] None of the three genes from isolates treated with DMI or QoI fungicides or in the untreated control group showed evidence of selection based on Tajima's *D* calculations.

Because gene-encoding regions are known to be less mutable than SSR loci, these regions were tested to determine their relative stability *in vitro* and resolution and sensitivity under field conditions. As expected based on microsatellite versus genic mutation rates, gene-encoding regions showed higher stability under fungicide-induced stress than SSR regions, which showed mutational differences at five of seven loci between progenitor and fungicide-exposed progeny. No mutational changes were observed in any of these progeny at the three gene-encoding regions. Untreated controls had no mutational changes at microsatellite loci, indicating that mutations were induced by the fungicide treatment and not likely the result of random mutation, as mutation rates at these loci are sufficiently small. This supports our hypothesis that fungicide stress does not influence gene regions to the same extent as it does SSR loci. Higher or longer fungicide stress, however, would likely increase potential for gene region mutation as fungicide-induced oxidative stress can overload the mismatch repair system, resulting in more uncorrected replication errors.[3,32] This would increase the likelihood of both genic mutations and microsatellite mutations.

For field isolates, we were surprised to find that the majority (72%) of the BB:TC isolate pairs from the same infection site had different MLGs between the blossom blight and twig canker isolate. One explanation for this phenomenon is that infection sites may contain more than one genotype in blossom or canker, and the variation came from observing different isolates. Comparing genotypes between infection sites, we detected two clonal genotypes using gene regions and no clonal isolates using SSRs. Although sequencing gene regions did not fully differentiate all 31 BB:TC isolates within infection sites, both methods were equally effective in differentiating isolates between infection sites and among orchards and trees. A lack of clonal genotypes was surprising because no apothecia were observed, and we expected that the asexual component of the *M. fructicola* life cycle would be most prevalent. However, the lack of clones is consistent with previous data and indicates that recombination is taking place despite an apparent lack of apothecia within stone fruit orchards.[12]

In a single season, both SSR loci and gene-encoding regions were able to distinguish isolates between different geographic locations, within the same tree and even within the same infection site for blossom-to-canker isolates. They are also both effective for studies that track the path of isolates over a generation. However, over longer periods of time with larger sample

sizes, the likelihood of mutational change in SSRs would increase as the SSR loci did in other studies,[16,19] potentially making gene-encoding regions a better marker choice.

Gene-encoding regions provide an alternative marker that may be helpful in verifying SSR data, as SSR genotyping can involve human error in allele calling. Although allele calling is mostly performed by software, human input is necessary to ensure that accurate allele detection occurs. This human involvement often results in genotyping inaccuracies.[33,34] Use of gene-encoding sequences along with SSRs or by themselves should help counteract this human error. Gene regions may also reduce the likelihood of error in genetic diversity calculations, as SSR loci are more likely to have alleles that are identical by size, but not by descent.[35]

Although there are definite advantages to using our gene region markers instead of SSRs, there are also limitations. Firstly, resolution of gene regions is slightly less than SSR marker resolution, which is preferred for genotyping closely related individuals. This idea is supported by the multirepeat differences in three loci between the two isolates that were not distinguished using gene regions, and these isolates were likely closely related spores that infected the same site. However, several studies suggest that the use of multiple marker types results in a more accurate characterization of the population structure because each marker type has different limitations, and using multiple marker types together can confirm that results are not artifacts of a particular marker type.[1] Based on the present study, a combination of SSR markers and the newly developed gene regions may be helpful to researchers examining population variation over time. Secondly, although our gene regions appear to be less vulnerable to mutation events, the field relevance of this is unknown as there are few studies characterizing mutagenesis under field conditions.

There are several ideal applications of these gene-region markers for studies on *M. fructicola*; for example, tracking fungicide-resistant isolates that are artificially introduced into an orchard over several seasons with fungicide applications to follow their spread and to determine their field fitness. These markers also enable additional field studies of fungicide or other stressor-induced mutagenesis over time. A proof of concept study using these markers was recently published, where markers were used to validate isolate identity when sampling the same infection site before and after application of a fungicide capable of causing mutagenesis *in vitro*.[21]

5 Conclusions

Overall, our new gene region markers for *M. fructicola* provide a more stable marker type that can be used to differentiate isolates within a field or even a single tree. Used alone, or combined with SSR regions, these new markers have the potential to give greater insight into the population dynamics of *M. fructicola*, with potential for future work to explore applicability of these markers to differentiate isolates of the closely related European brown rot species *M. laxa and M. fructigena*.

Acknowledgments — This project was supported by Technical Contribution No. 6401 of the Clemson University Experiment Station and by NIFA/USDA under project number SC-1700501. We thank Jeff Hopkins, Sam Hudson, Ryan McNeill and the rest of the staff at the Musser and UGA Horticulture Research Farms for assistance in fungicide applications and orchard maintenance. We also thank Dr Hideo Ishii, Dr Mengjun Hu, Fei Fan, Zhen Fan, Simon Li, Xi Song and Jing-Hui Yang for their help with isolate collection. Special thanks also go to Xingpeng 'Simon' Li for sharing transcriptome assembly data with us.

Supporting Information — Supporting information follows the References.

References

1 Arnaud-Haond S, Alberto F, Teixeira S, Procaccini G, Serrao EA, Duarte CM *et al.*, Assessing genetic diversity in clonal organisms: low diversity or low resolution? Combining power and cost efficiency in selecting markers. *J Hered* **96**:434–440 (2005).

2 Kohn LM, Schaffer MR, Anderson JB and Grünwald NJ, Marker stability throughout 400 days of *in vitro* hyphal growth in the filamentous ascomycete, *Sclerotinia sclerotiorum*. *Fungal Genet Biol* **45**:613–617 (2008).

3 Evans MD and Cooke MS, Factors contributing to the outcome of oxidative damage to nucleic acids. *Bioessays* **26**:533–542 (2004).

4 Ross KG, Shoemaker DD, Krieger MJ, DeHeer CJ and Keller L, Assessing genetic structure with multiple classes of molecular markers: a case study involving the introduced fire ant *Solenopsis invicta*. *Mol Biol Evolution* **16**:525–543 (1999).

5 Scherm H and Emery KM, Vegetative compatibility in populations of *Monilinia fructicola* from Georgia peach orchards, in *V International Peach Symposium, Davis, CA, 2001*, ed. by Scott Johnson R and Chrisosto CH. Acta Horticulturae Books No. 592, International Society for Horticultural Science, Leuven, Belgium (2002).

6 Villarino M, Larena I, Martinez F, Melgarejo P and De Cal A, Analysis of genetic diversity in *Monilinia fructicola* from the Ebro Valley in Spain using ISSR and RAPD markers. *Eur J Plant Pathol* **132**:511–524 (2012).

7 McDonald BA, The population genetics of fungi: tools and techniques. *Phytopathology* **87**:448–453 (1997).

8 Sonoda RM, Ogawa JM, Esser TE and Manji BT, Mycelial interaction zones among single ascospore isolates of *Monilinia fructicola*. *Mycologia* **74**:681–683 (1982).

9 Emery KM, Scherm H and Savelle AT, Assessment of interactions between components of fungicide mixtures against *Monilinia fructicola*. *Crop Prot* **21**:41–47 (2002).

10 Fan JY, Guo LY, Xu JP, Luo Y and Michailides TJ, Genetic diversity of populations of *Monilinia fructicol*a (fungi, Ascomycota, Helotiales) from China. *J Eukaryot Microbiol* **57**:206–212 (2010).

11 Everhart SE, Askew A, Seymour L, Glenn TC and Scherm H, Spatial patterns of brown rot epidemics and development of microsatellite markers for analyzing fine-scale genetic structure of *Monilinia fructicola* populations within peach tree canopies. *Plant Hlth Prog* DOI:10.1094/PHP-2012-0723-04-RS (2012).

12 Everhart S and Scherm H, Fine-scale genetic structure of *Monilinia fructicola* during brown rot epidemics within individual peach tree canopies. *Phytopathology* **105**:542–549 (2015).

13 Selkoe KA and Toonen RJ, Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol Lett* **9**:615–629 (2006).

14 Lynch M, The origins of eukaryotic gene structure. *Mol Biol Evol* **23**:450–468 (2006).

15 Jackson AL and Loeb LA, Microsatellite instability induced by hydrogen peroxide in *Escherichia coli*. *Mutat Res Fund Mol Mech Mut* **447**:187–198 (2000).

16 Amaradasa BS and Everhart SE, Effects of sublethal fungicides on mutation rates and genomic variation in fungal plant pathogen, *Sclerotinia sclerotiorum*. *PLoS ONE* **11**:e0168079 (2016).

17 Hasanuzzaman M, Nahar K, Fujita M, Vahdati K and Leslie C, *Extreme Temperature Responses, Oxidative Stress and Antioxidant Defense in Plants*. InTech, Rijeka, Croatia (2013).
18 Troncoso-Rojas R, Báez-Flores ME, Pryor B, García HS and Tiznado-Hernández M, Inter simple sequence repeat polymorphism in *Alternaria* genomic DNA exposed to lethal concentrations of isothiocyanates. *Afr J Microbiol Res* **7**:838–852 (2013).
19 Schnabel G, Chen F, Everhart SE, Bridges WC and Liu XL, Studies on sensitivity reduction in solo and mixture treatments and fungicide-induced mutagenesis in *Monilinia fructicola*, ed. by Dehne HW, Deising HB, Gisi U, Fraaije B, Hermann D, Mehl A. "Modern Fungicides and Antifungal Compounds", Vol. VII, Deutsche Phytomedizinische Gesellschaft, Braunschweig, ISBN: 978-3-941261-13-6, pp 263–268 (2014).
20 Chen F, Everhart SE, Bryson PK, Luo C, Song X, Liu X *et al.*, Fungicide-induced transposon movement in *Monilinia fructicola*. *Fungal Genet Biol* **85**:38–44 (2015).
21 Dowling ME, Bryson PK, Boatwright HG, Wilson JR, Fan Z, Everhart SE *et al.*, Effect of fungicide applications on *Monilinia fructicola* population diversity and transposon movement. *Phytopathology* **106**:1504–1512 (2016).
22 Li B and Dewey CN, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf* **12**:1 (2011).
23 Thorvaldsdottir H, Robinson JT and Mesirov JP, Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinf* **14**:178–192 (2013).
24 Horton D, Brannen P, Bellinger B, Lockwood D and Ritchie D, *Southeastern Peach, Nectarine and Plum Pest Management and Culture Guide*. Bulletin 1171, University of Georgia, Athens, GA (2013).
25 Chi M, Park S and Lee Y, A quick and safe method for fungal DNA extraction. *Plant Pathol J* **25**:108–111 (2009).
26 Kamvar ZN, Tabima JF and Grünwald NJ, Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *Peer J* **2**:e281 (2014).
27 Tajima F, Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**:585–595 (1989).
28 Librado P and Rozas J, DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**:1451–1452 (2009).
29 Leroch M, Plesken C, Weber RW, Kauff F, Scalliet G and Hahn M, Gray mold populations in German strawberry fields are resistant to multiple fungicides and dominated by a novel clade closely related to *Botrytis cinerea*. *Appl Environ Microbiol* **79**:159–167 (2013).
30 Chen W, Lee MK, Jefcoate C, Kim SC, Chen F and Yu JH, Fungal cytochrome p450 monooxygenases: their distribution, structure, functions, family expansion, and evolutionary origin. *Genome Biol Evol* **6**:1620–1634 (2014).
31 Dunkel N, Blass J, Rogers PD and Morschhauser J, Mutations in the multi-drug resistance regulator MRR1, followed by loss of heterozygosity, are the main cause of MDR1 overexpression in fluconazole-resistant *Candida albicans* strains. *Mol Microbiol* **69**:827–840 (2008).
32 Jackson AL, Chen R and Loeb LA, Induction of microsatellite instability by oxidative DNA damage. *Nat Acad Sci Proc* **95**:12 468–12 473 (1998).
33 Hoffman JI and Amos W, Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Mol Ecol* **14**:599–612 (2004).
34 Pompanon F, Bonin A, Bellemain E and Taberlet P, Genotyping errors: causes, consequences and solutions. *Nat Rev Genet* **6**:847–846 (2005).
35 Taylor JW and Fisher MC, Fungal multilocus sequence typing – it's not just for bacteria. *Curr Opin Microbiol* **6**:351–356 (2003).

Supplemental Table 1 – Genotypes of isolates (62) in the present study that were matching BB:TC pairs

			SSR[a] Loci							Gene Regions[b]		
Site/Isolate ID	**Location**	**Symptom**	**SED**	**SEF**	**SEI**	**SEL**	**SEN**	**SEP**	**SEQ**	***mrr1***	***DHFR***	***MfCYP01***
10.1.8.bb	Musser Farm	BB	131	139	114	137	227	265	134	26	1	3
10.1.8.tc	Musser Farm	TC	131	139	114	137	227	265	134	26	1	3
10.2.2.bb	Musser Farm	BB	145	117	114	147	223	254	134	30	7	2
10.2.2.tc	Musser Farm	TC	145	117	114	147	223	254	134	30	7	2
10.3.2.bb	Musser Farm	BB	145	117	110	147	227	254	132	12	3	3
10.3.2.tc	Musser Farm	TC	145	117	110	147	227	254	132	12	3	3
10.5.17.bb	Musser Farm	BB	135	123	114	147	227	257	138	8	4	2
10.5.17.tc	Musser Farm	TC	135	123	114	147	227	257	138	8	4	2
10.5.7.bb	Musser Farm	BB	127	126	114	143	227	254	140	12	4	3
10.5.7.tc	Musser Farm	TC	127	126	114	143	227	254	140	12	4	3
3.1.6.bb	Musser Farm	BB	131	126	110	137	227	250	136	20	7	17
3.1.6.tc	Musser Farm	TC	131	126	110	137	227	250	136	20	7	17
3.1.8.bb	Musser Farm	BB	145	131	114	147	227	254	134	2	11	1
3.1.8**.tc**	Musser Farm	TC	145	131	114	147	227	254	134	2	11	3
3.2.17.bb	Musser Farm	BB	131	117	114	137	231	247	138	32	7	3
3.2.17.tc	Musser Farm	TC	131	117	114	137	231	247	138	32	7	3
3.2.7.bb	Musser Farm	BB	141	131	114	143	231	247	134	24	21	13
3.2.7.tc	Musser Farm	TC	141	131	114	143	231	247	134	24	21	13
3.2.8.bb	Musser Farm	BB	131	117	104	143	227	254	132	27	4	8
3.2.8.tc	Musser Farm	TC	131	117	104	143	227	254	132	27	4	8
3.4.1.bb	Musser Farm	BB	131	131	114	147	223	247	132	9	20	2
3.4.1.tc	Musser Farm	TC	131	131	114	147	223	247	132	9	20	2
4.1.4.bb	Musser Farm	BB	135	144	114	137	231	254	130	14	12	16
4.1.4.tc	Musser Farm	TC	135	144	114	137	231	254	130	14	12	16
4.5.11.bb	Musser Farm	BB	131	126	104	137	231	231	138	20	4	15
4.5.11.tc	Musser Farm	TC	131	126	104	137	231	231	138	20	4	15
6.1.8.bb	Musser Farm	BB	141	139	113	147	223	254	138	9	11	1
6.1.8.tc	Musser Farm	TC	141	139	113	147	223	254	138	9	11	1
6.2.13.bb	Musser Farm	BB	133	135	114	143	231	254	146	29	4	2
6.2.13.tc	Musser Farm	TC	133	135	114	143	231	254	146	29	4	2
6.3.18.bb	Musser Farm	BB	131	117	114	137	227	254	140	13	22	1

6.3.18.tc	Musser Farm	TC	131	117	114	137	227	254	140	13	22	1
6.4.4.bb	Musser Farm	BB	141	123	102	143	231	231	134	28	23	4
6.4.4.tc	Musser Farm	TC	141	123	102	143	231	231	134	28	23	4
7.1.1.bb	Musser Farm	BB	131	117	110	143	227	254	138	12	4	4
7.1.1.tc	Musser Farm	TC	131	117	110	143	227	254	138	12	4	4
7.2.17.bb	Musser Farm	BB	131	117	98	137	227	254	138	4	20	3
7.2.17.tc	Musser Farm	TC	131	117	98	137	227	254	138	4	20	3
7.3.12.bb	Musser Farm	BB	131	117	114	143	235	250	140	8	4	8
7.3.12.tc	Musser Farm	TC	131	117	114	143	235	250	140	8	4	8
7.5.4.bb	Musser Farm	BB	131	127	104	137	227	250	138	24	7	7
7.5.4.tc	Musser Farm	TC	131	127	104	137	227	250	138	24	7	7
9.2.18.bb	Musser Farm	BB	131	117	114	143	231	254	134	9	4	1
9.2.18.tc	Musser Farm	TC	131	117	114	143	231	254	134	9	4	1
9.5.3.bb	Musser Farm	BB	131	131	114	137	223	254	136	15	11	8
9.5.3.tc	Musser Farm	TC	131	131	114	137	223	254	136	15	11	8
9.5.6.bb	Musser Farm	BB	131	144	104	143	235	231	136	15	4	7
9.5.6.tc	Musser Farm	TC	131	144	104	143	235	231	136	15	4	7
9.5.7.bb	Musser Farm	BB	131	144	114	137	231	254	138	29	4	2
9.5.7.tc	Musser Farm	TC	131	144	114	137	231	254	138	29	4	2
Azo13.bb	UGA Horticulture Farm	BB	106	117	114	137	223	254	130	31	30	1
Azo13.tc	UGA Horticulture Farm	TC	106	117	114	137	223	254	130	31	30	1
Azo82.bb	UGA Horticulture Farm	BB	127	149	110	143	227	254	125	22	15	3
Azo82.tc	UGA Horticulture Farm	TC	127	149	110	143	227	254	125	22	15	3
Azo93.bb	UGA Horticulture Farm	BB	127	127	114	137	227	254	138	15	4	3
Azo93.tc	UGA Horticulture Farm	TC	127	127	114	137	227	254	136	15	4	3
Con176.bb	UGA Horticulture Farm	BB	141	135	117	143	223	254	136	10	1	11
Con176.tc	UGA Horticulture Farm	TC	141	135	117	143	223	257	134	10	1	11
Con66.bb	UGA Horticulture Farm	BB	106	131	98	137	231	257	134	25	7	14
Con66.tc	UGA Horticulture Farm	TC	106	131	98	137	231	250	134	25	7	14
Con91.bb	UGA Horticulture Farm	BB	106	126	114	143	227	250	134	3	16	3
Con91.tc	UGA Horticulture Farm	TC	106	126	114	143	227	250	134	3	16	3

[a]Simple sequence repeat

[b]Numbers were assigned to represent each different genotype